NAG Fortran Library Routine Document

G01DBF

Note: before using this routine, please read the Users' Note for your implementation to check the interpretation of **bold italicised** terms and other implementation-dependent details.

1 Purpose

G01DBF calculates an approximation to the set of Normal Scores, i.e., the expected values of an ordered set of independent observations from a Normal distribution with mean 0.0 and standard deviation 1.0.

2 Specification

```
SUBROUTINE G01DBF(N, PP, IFAIL)
INTEGER        N, IFAIL
real           PP(N)
```

3 Description

This routine is an adaptation of the Applied Statistics Algorithm AS 177.3, see Royston (1982). If the user is particularly concerned with the accuracy with which G01DBF computes the expected values of the order statistics (see Section 7), then G01DAF which is more accurate should be used instead at a cost of increased storage and computing time.

Let $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ be the order statistics from a random sample of size n from the standard Normal distribution. Defining

$$P_{r,n} = \Phi(-E(x_{(r)}))$$

and

$$Q_{r,n} = \frac{r - \epsilon}{n + \gamma}, \quad r = 1, 2, \ldots, n$$

where $E(x_{(r)})$ is the expected value of $x_{(r)}$, the current routine approximates the Normal upper tail area corresponding to $E(x_{(r)})$ as,

$$\tilde{P}_{r,n} = Q_{r,n} + \frac{\delta_1}{n} Q_{r,n}^{\lambda} + \frac{\delta_2}{n} Q_{r,n}^{2\lambda} - C_{r,n}.$$

Estimates of ϵ, γ, δ_1, δ_2 and λ are obtained for $r = 1, 2, 3$, and $r \geq 4$. A small correction $C_{r,n}$ to $\tilde{P}_{r,n}$ is necessary when $r \leq 7$ and $n \leq 20$.

The approximation to $E(X_{(r)})$ is thus given by

$$E(x_{(r)}) = -\Phi^{-1}(\tilde{P}_{r,n}), \quad r = 1, 2, \ldots, n.$$

Values of the inverse Normal probability integral Φ^{-1} are obtained from G01FAF.

4 References

Royston J P (1982) Algorithm AS 177: Expected Normal order statistics (Exact and Approximate) *Appl. Statist.* **31** 161–165

5 Parameters

1: N – INTEGER *Input*

On entry: the size of the sample, n.

Constraint: $N \geq 1$.

2: PP(N) – ***real*** array *Output*

On exit: the Normal scores. $PP(i)$ contains the value $E(x_{(i)})$, for $i = 1, 2, \ldots, n$.

3: IFAIL – INTEGER *Input/Output*

On entry: IFAIL must be set to 0, -1 or 1. Users who are unfamiliar with this parameter should refer to Chapter P01 for details.

On exit: IFAIL $= 0$ unless the routine detects an error (see Section 6).

For environments where it might be inappropriate to halt program execution when an error is detected, the value -1 or 1 is recommended. If the output of error messages is undesirable, then the value 1 is recommended. Otherwise, for users not familiar with this parameter the recommended value is 0. **When the value -1 or 1 is used it is essential to test the value of IFAIL on exit.**

6 Error Indicators and Warnings

If on entry IFAIL $= 0$ or -1, explanatory error messages are output on the current error message unit (as defined by X04AAF).

Errors or warnings detected by the routine:

IFAIL $= 1$

On entry, $N < 1$.

7 Accuracy

For $n \leq 2000$, the maximum error is 0.0001, but the routine is usually accurate to 5 or 6 decimal places. For n up to 5000, comparison with the exact scores calculated by G01DAF shows that the maximum error is 0.001.

8 Further Comments

The time taken by the routine is proportional to n.

9 Example

A program to calculate the expected values of the order statistics for a sample of size 10.

9.1   Program Text

Note: the listing of the example program presented below uses ***bold italicised*** terms to denote precision-dependent details. Please read the Users' Note for your implementation to check the interpretation of these terms. As explained in the Essential Introduction to this manual, the results produced may not be identical for all implementations.

```
*      G01DBF Example Program Text
*      Mark 14 Revised.  NAG Copyright 1989.
*      .. Parameters ..
       INTEGER          N
       PARAMETER        (N=10)
       INTEGER          NOUT
       PARAMETER        (NOUT=6)
*      .. Local Scalars ..
       INTEGER          I, IFAIL
*      .. Local Arrays ..
       real             PP(N)
*      .. External Subroutines ..
       EXTERNAL         G01DBF
*      .. Executable Statements ..
       WRITE (NOUT,*) 'G01DBF Example Program Results'
       IFAIL = 0
*
       CALL G01DBF(N,PP,IFAIL)
*
       WRITE (NOUT,*)
       WRITE (NOUT,99999) 'Sample size = ', N
       WRITE (NOUT,*) 'Normal scores'
       WRITE (NOUT,99998) (PP(I),I=1,N)
       STOP
*
99999 FORMAT (1X,A,I2)
99998 FORMAT (10X,5F12.4)
       END
```

9.2   Program Data

None.

9.3   Program Results

```
 G01DBF Example Program Results

 Sample size = 10
 Normal scores
              -1.5388      -1.0014      -0.6561      -0.3757      -0.1227
               0.1227       0.3757       0.6561       1.0014       1.5388
```